

GTG	TTC	AGC	TCC	GAG	GCC	GCC	TGG	CAG	TGT	GTG	AGT	GAG	GCG	CTG	CAG	ATC	CTC	GGG	GGC	1209
L	G	Y	T	R	D	Y	P	Y	E	R	I	L	R	D	T	R	I	L	L	423
TTG	GCG	TAC	ACA	AGG	GAC	TAT	CCG	TAC	GAG	CGC	ATA	CTG	GAC	ACC	CGC	ATC	CTC	CTC	1269	
I	F	E	G	T	N	E	I	L	R	M	Y	I	A	L	T	G	L	Q	H	443
ATC	TTC	GAG	GGA	ACC	AAT	GAG	ATT	CTC	CGG	ATG	TAC	ATC	GCC	CTG	ACG	GGT	CTG	CAG	CAT	1329
A	G	R	I	L	T	T	R	I	H	E	L	K	Q	A	K	V	S	T	V	463
GCC	GCG	CGC	ATC	CTG	ACT	ACC	AGG	ATC	CAT	GAG	CTT	AAA	CAG	GCC	AAA	GTG	AGC	ACA	GTC	1389
M	D	T	V	G	R	R	L	R	D	S	L	G	R	T	V	D	L	G	L	483
ATG	GAT	ACC	GTT	GGC	CGG	AGG	CTT	CGG	GAC	TCC	CTG	GGC	CGA	ACT	GTG	GAC	CTG	GGG	CTG	1449
T	G	N	H	G	V	V	H	P	S	L	A	D	S	A	N	K	F	E	E	503
ACA	GGC	AAC	CAT	GGA	GTT	GTG	CAC	CCC	AGT	CTT	GCG	GAC	AGT	GCC	AAC	AAG	TTT	GAG	GAG	1509
N	T	Y	C	F	G	R	T	V	E	T	L	L	L	R	F	G	K	T	I	523
AAC	ACC	TAC	TGC	TTC	GGC	CGG	ACC	GTG	GAG	ACA	CTG	CTG	CTC	CGC	TTT	GGC	AAG	ACC	ATC	1569
M	E	E	Q	L	V	L	K	R	V	A	N	I	L	I	N	L	Y	G	M	543
ATG	GAG	GAG	CAG	CTG	GTA	CTG	AAG	CGG	GTG	GCC	AAC	ATC	CTC	ATC	AAC	CTG	TAT	GGC	ATG	1629
T	A	V	L	S	R	A	S	R	S	I	R	I	G	L	R	N	H	D	H	563
ACG	GCC	GTG	CTG	TGC	CGG	GCC	AGC	CGC	TCC	ATC	CGC	ATT	GGG	CTC	CGC	AAC	CAC	GAC	CAC	1689
E	V	L	L	A	N	T	F	C	V	E	A	Y	L	Q	N	L	F	S	L	583
GAG	GTT	CTC	TTG	GCC	AAC	ACC	TTC	TGC	GTG	GAA	GCT	TAC	TTG	CAG	AAT	CTC	TTC	AGC	CTC	1749
S	Q	L	D	K	Y	A	P	E	N	L	D	E	Q	I	K	K	V	S	Q	603
TCT	CAG	CTG	GAC	AAG	TAT	GCT	CCA	GAA	AAC	CTA	GAT	GAG	CAG	ATT	AAG	AAA	GTG	TCC	CAG	1809
Q	I	L	E	K	R	A	Y	I	C	A	H	P	L	D	R	T	C	*		622
CAG	ATC	CTT	GAG	AAG	CGA	GCC	TAT	ATC	TGT	GCC	CAC	CCT	CTG	GAC	AGG	ACA	TGC	TGA		1866
GGC	AGG	GAC	AGT	GTCCCT	GCT	TACC	GCCCG	CCCC	TACCC	ATGGCCG	TTGCT	GGAT	GACT	GTT	TACT	CTTTTT	T	CAGAA		
GCTG	TTGGG	ATTAT	CACAG	GTTA	AGCCT	TTTGT	TCCCG	CTCTG	CACCT	GAAAGG	TTTGT	CGCCT	GGCCT	TGGG	AGAG	AGCCT	C			
TTCC	AGGTTT	TGACCT	GCAGG	CAGT	GCTCT	CTA	CAGG	ACCAT	CACAG	CTTCT	GAACT	GAGCCG	AGAG	AGAGA	ATGGA					
ATTG	CTG	ACCC	TGGA	ACTGG	CGGGT	TATCT	GGT	CATTG	AGG	GAGAC	ACCAT	AGT	TGGA	AACT	TGGG	GCTTAT	GCTG	CTGCC		
TCC	AGGGT	GTG	AGGT	GGTGGG	GACCT	GTG	TCA	GGT	GTG	GATAG	CCATT	TTCT	GCTCA	ACCAC	ATTCT	CTA	AGAA	ACA		
GCTT	GAA	AGCT	CTG	CTGGG	CATT	CATT	TAA	ACT	TAGA	AGCAG	AGGCA	CTTAA	AAAC	ATG	TAC	CAGGA	ACCATT	TAA	CAA	
AGA	ATATA	AAAA	TG	CA	CAAT	CTG	TG	TACT	GT	TAAAA	AAAAAAAA									

FIGURE 1B

Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM

hmmpfam - search a single seq against HMM database

HMMER 2.1.1 (Dec 1998)

Copyright (C) 1992-1998 Washington University School of Medicine

HMMER is freely distributed under the GNU General Public License (GPL).

HMM file: /prod/ddm/seqanal/PFAM/pfam6.4/Pfam

Sequence file: /prod/ddm/wspace/orfanal/oa-script.26629.seq

Query: 62112

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
-----	-----	-----	-----	-----
Acyl-CoA_dh_M 1	Acyl-CoA dehydrogenase, middle domain	153.0	3.5e-42	
Acyl-CoA_dh 1	Acyl-CoA dehydrogenase, C-terminal dom	152.1	9.6e-42	
Acyl-CoA_dh_N 1	Acyl-CoA dehydrogenase, N-terminal dom	73.7	4.2e-19	
Polysac_deacet 1	Polysaccharide deacetylase	-43.7	1.8	

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
-----	-----	-----	-----	-----	-----	-----	-----
Acyl-CoA_dh_N	1/1	85	177	..	29 132 .]	73.7	4.2e-19
Acyl-CoA_dh_M	1/1	179	286	..	1 106 []	153.0	3.5e-42
Acyl-CoA_dh	1/1	290	441	..	1 156 []	152.1	9.6e-42
Polysac_deacet	1/1	432	580	..	1 150 []	-43.7	1.8

Alignments of top-scoring domains:

Acyl-CoA_dh_N: domain 1 of 1, from 85 to 177: score 73.7, E = 4.2e-19	
*->RRVDKsgefPalrelrikaLgqlGllginvPEeyGGaGad..ylarFm	
R++D++g+ P e +++L lG1+g+ vPEyGG+G +++ ++	
62112 85 RKIDQEGKIP--DETLEKLKSLGLGLQVPEEYGGLGFSntMYS--- 126	
LHAQVaalviEElarvcAastgvilsvhssLggnpilkfGseEQKkylPq	
+ E+ ++s v+l++h ++g+ +i+ +G+eEQK+kyLp+	
62112 127 -----RLGETIISMDGSITVTTLAAHQAILGLKGIILAGTEQKAKYLPK 168	
ltsGdliga<-*	
l+sG++i+a	
62112 169 LASGEHIAA 177	
Acyl-CoA_dh_M: domain 1 of 1, from 179 to 286: score 153.0, E = 3.5e-42	
*->ALTEPgAGSDvgsSlkTtAekkEGD..dyiLNGsKmWITNGgqAdwyi	
+lTEP +GSD+S++ +A+ d+++yilNGsK+WITNGg A++++	
62112 179 CLTEPASGSDAASIRSRATLS-EDkKhYILNGSKVWITNGGLANIFT 224	
VlAvT...DpakkvpgkkgitaFlVekdtpGfsaGkKedKLGLRgsdTeE	
V+A+T+ D + + k itaF+Ve+d+ G++ Gk+edKLG+RgS+TcE	
62112 225 VFAKTeVvDSDG--SVKDKITAFIVERDFGGVTNGKPEDKLIRGSNTCE 272	
LiFEDvrvPesniL<-*	
+ FE+ ++P +nIL	
62112 273 VHFENTKIPVENIL 286	

Acyl-CoA_dh: domain 1 of 1, from 290 to 441: score 152.1, E = 9.6e-42

*->GkGFKyamkeLdmeRlviAaqalGlaggaldeAinYakgRkgFGkpl	
G+GFK+am+ L+ +R+ +++ Gl+ + + + +Ya RkgF k+l	
62112 290 GDGFKVAMNILNSGRFSMGSVVAGLLKRLIEMTAeyACTRkQPNKRL 336	

FIGURE 2A

```

        adfQliQfkLAdMatkLEaarllvYraAwladr.GedAKEALptskeaaam
        +f liQ+k+A Ma k ++ +++Y +A d++G + ++s eaam
62112 337 SEFGLIQEKFALMAQKAYVMESMTYLTAGMLDQpGFP-----DCSIEAAM 381

        AKlfaseaAmgvatdAvQilGGvGYtkdyPveRfyRDAkitqIVEGTsEI
        +K f+seaA q +++A+QilGG GYt dyP eR +RD +I I EGT+EI
62112 382 VKVFSSEAAWQCvSEALQILGGLGYTRDYPYERILRDTRILLIFEGTNEI 431

        qrlvIaRall<-*
        r Ia + l
62112 432 LRMYIALTGL 441

Polysac_deacet: domain 1 of 1, from 432 to 580: score -43.7, E = 1.8
        *->ddksvylTFDDGpNaApayTprILDvLkhhkvkATFFviGsnvkdnP
        +++++LT + ++ + T+r+ + Lk+ kv + G++ +d
62112 432 LRMYIALTGLQHAG--RILTTRI-HELKQAKVSTVMDTVGRRLRD-- 473

        dlarrivkeGHeignNhtwsHPdlt.....t1
        + r v+ G gNH+ HP l+++ ++ +++++ + + + +
62112 474 -SLGRTVDLG-LTGNHGVVHPSLAdsankfeentycfgrtvettlllrfgK 521

        taeqirdeiartneaiigatggatptlfrpPYGewsetvlsasaklGlaa
        t +++ + r++++i+++g t++l R+ s+s ++Gl+
62112 522 TIMEERQLVKRVANILINLYG-MTAVLSRA-----SRSIRIGLRN 560

        vlWdvDprDwsvragadaivdavlga<-*
        + D v ++ v a+lg+
62112 561 H-----DHEVLLANTFCVEAYLQNL 580

```

FIGURE 2B

Protein Family / Domain Matches, HMMer version 2

Searching for complete domains in PFAM

hmmpfam - search a single seq against HMM database

HMMER 2.1.1 (Dec 1998)

Copyright (C) 1992-1998 Washington University School of Medicine

HMMER is freely distributed under the GNU General Public License (GPL).

HMM file:

/prod/dam/seqanal/PFAM/pfam5.4/Pfam

Sequence file:

/prod/dam/wspace/orfanal/oa-script.17193.seq

Query: 62112

Scores for sequence family classification (score includes all domains):

Model	Description	Score	E-value	N
Acyl-CoA_dh	Acyl-CoA dehydrogenase	399.8	1.8e-116	1
Polysac_deacet	Polysaccharide deacetylase	-43.7	1	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
Acyl-CoA_dh	1/1	85	438	..	29 394 .1	399.8	1.8e-116
Polysac_deacet	1/1	432	580	..	1 150 []	-43.7	1

Alignments of top-scoring domains:

Acyl-CoA_dh: domain 1 of 1, from 85 to 438: score 399.8, E = 1.8e-116

*->RRvDksgefPlrelikaLgklGllginvPEeyGGsGad..ylaRfMl

R++D+g+ P e ++L 1G1+g+ vPEeyGG+G +++ ++

62112 85 RKIDQBGKIP-DETLKLSLGLFGLQVPEEYGGLSntMYS---- 126

HAQVaalviEElarvcAstgvllsvhssLgqnpilrfGseEQkkkyLpql

+ E+ ++s v+l++h ++g+ i+ +G+eEQk+kyLp+l

62112 127 -----RLGEIISMDGSITVTTLAAHQAIQLKGIILAGTEEQKAKYLPKL 169

tsGdligafALTEPgAGSDvgSikTtAekkeEGD..dyiLNGsKmWITNGg

FIGURE 2C

```

62112 170 +sG++i+af+lTEP +GSD++Si+ +A+ d+++yILNGsK+WITNGG
ASGEHIAAFCLTEPASGSDAASIRSRTLs-EDKkHYILNGSKVWITNGG 218

qAdwyiVlAVT...DpakkvpgkkgitaFlVekdtpGfsiGkKedKLGIR
A++++V+A+T+ D + + k itaF+Ve+d+ G++ Gk+edKLG+R
62112 219 LANIFTVFAKTeVvSDG--SVKDKITAFIVERDFGGVtNGKPEDKLGIR 266

gSdTcELiFEDvrvPesniLGeEGeGfKyamKtLdmeRlgiAaqalGlaq
gS+Tce+ FE+ ++P +niLGe G+GfK+aM+ L+ +R+ +++ G++
62112 267 GSNTCEVHFENTKIPVENILGEVGDGfKVAMNILNSGRFSMGSVVAGLLK 316

gAldeAinYAkqRkqFGkplaeFQliQfKLAdMatkLEaarllyYraAwl
+ ++ +++YA RkqF k+l ef liQ+k+A MA k ++ +++Y +A
62112 317 RLIEMTAeyACTRKQFNKRLSEfGLIQEKfALMAQAYVMESMTYLTAGM 366

adr.GedAKEALptskeAAMAKlfAseiAmkvatdAvQilGGvGYtkdyP
d++G + ++s eAAM+K f+se+A + +++A+QilGG Gyt dyP
62112 367 LDQpGFP----DCSIEAMVKVFSSEAAWQCVSEALQILGGLGYTRDYP 411

veRfyRDAkitqIYEGTsEIQrlvIaR<--*
eR +RD +i I EGT+EI r Ia
62112 412 YERILRDTRILLIFEgTNEILRMylAL 438

Polysac_deacet: domain 1 of 1, from 432 to 580: score -43.7, E = 1
*->ddksvylTFDDGPnAApayTprlLDvLkKhkvkATFFviGsnvkdnp
++++LT + ++ + T+r+ + Lk+ kv + G++ +d
62112 432 LRMylALTGLQHAG--RILtTRI--HELKQAKVSTVMDTVGRRLRD-- 473

dlarriVkeGHeigNhtwsHPdlT.....t1
+ r v+ G gNH+ HP l+++ ++ +++ + + +
62112 474 -SLGRTVDLG-LTGNGHVvHPSLAdsankfeentycfgrtvetlllrfG 521

taeqirdeiertneaiiqatggatptlfrPyGwsetvlsasaklGltA
t +++ + r++++i+++g t+l R+ s+s ++Gl+
62112 522 TIMEQLVLKRVANILINLYG-MTAVLSRA-----SRsIRIGLRN 560

vLWdvDprDwsvradaivdavlqaa<--*
+ D v ++ v a+lq+
62112 561 H-----DHEVLLANTFCVEAYLQNL 580

```

FIGURE 2D

TxP expression: 30K array

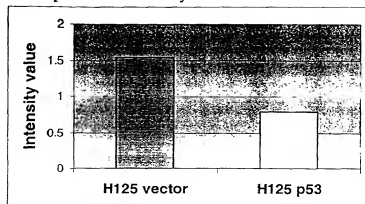


FIGURE 3A

Taqman expression

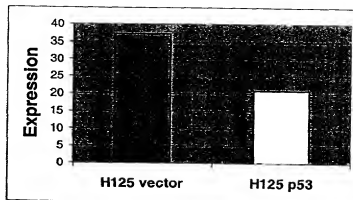


FIGURE 3B

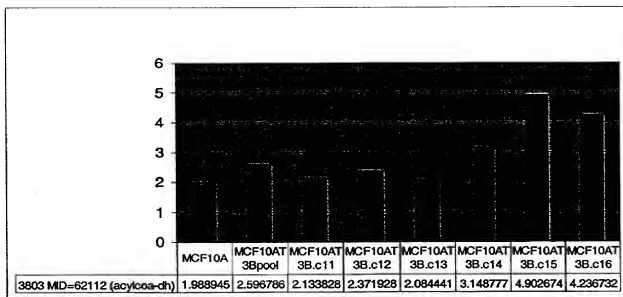


FIGURE 4

101680" 92251660

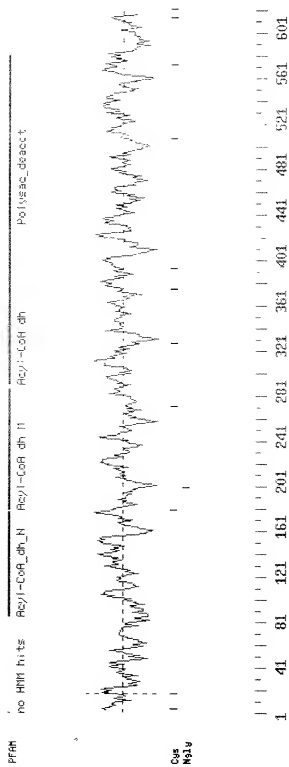


FIGURE 5